example, general file names instead of URLs may be used in a distributed file sharing system with

compressed Bloom filters.

5   WHAT IS CLAIMED IS:

1   1. A distributed, compressed bloom filter Web server providing reduced probabilities of false

2   positives, comprising:

3   a plurality of cache servers each having a cache memory and a cache processor coupled to

4   the memory that is operative (1) to represent Web objects stored in its cache memory as a

5   compressed Bloom filter data array having a preselected number of hash functions and a preselected

6   array size which have been chosen to minimize the rate of false positives for a preselected

7   transmission compression size, (2) to compress the Bloom filter data array to said transmission

8   compression size, and (3) to periodically disseminate the compressed Bloom filter data array to

9   neighboring servers when there is a change in its stored Web objects.

1   2. The distributed, compressed Bloom filter Web server providing reduced probabilities of false

2   positives of claim 1 , wherein the Bloom filter data array size is made as large as possible for a given

3   cache memory size and the number of hash functions is chosen to minimize the rate of false positives

4   for a preselected transmission compression size.

1   3. The distributed, compressed Bloom filter Web server providing reduced probabilities of false

2   positives of claim 1, wherein arithmetic coding is employed to compress the Bloom filter data array

3   to said transmission compression size.

1   4. The distributed, compressed Bloom filter Web server providing reduced probabilities of false

2   positives of claim 1, wherein said cache processor is further operative to store in its cache memory

3   at least one other decompressed Bloom filter data array each representative of the Web objects of

4   another, at least one neighboring Web server.

1   5. A method reducing false positives in a network having distributed Web servers each storing

2   information in cache memory as a Bloom filter data array representative of the information in its

3   cache memory and broadcasting that data array to other Web servers periodically, comprising:

4   1) fixing a compression size for transmission of the Bloom filter data array at each Web

5   server;

6   (2) choosing the number of hash functions and the array size so that the Bloom filter data

7   array minimizes the rate of false positives when it is compressed down to the fixed compression size;

8   and

9   (3) broadcasting the compressed Bloom filter data array of the fixed compression size to

10   neighboring Web servers whenever there is a change in the contents of its cache memory.

1   6. A distributed computer network, comprising:

2   a plurality of periodically intercommunicating distributed network nodes;

3   each node including a cache memory and a processor coupled to the cache memory operative

4   to (1) represent its memory contents as a compressed Bloom filter data structure having a preselected

22

5   number of hash functions and a preselected array size which have been chosen for a target

6   compression size to optimize at least one of the rate of false positives of the Bloom filter

7   representing the memory contents and the computational requirements of the preselected number of

8   hash functions, to (2) compress the Bloom filter data structure to the target compression size using

9   a predetermined compression algorithm, and to (3) broadcast the compressed Bloom filter data

10   structure to at least one other node whenever the contents of its cache memory has changed.

1   7. The distributed computer network of claim 1, wherein said nodes are Web proxy servers.

1   8. The distributed computer network of claim 1, wherein said nodes are mobile or stationary agents

2   in a network of mobile nodes, and the Web objects correspond to agent locations.

1   9. The distributed computer network of claim 1, wherein said predetermined compression algorithm

2   is arithmetic coding.

1   10. A method employing compressed Bloom filters for storing and transmitting data in a distributed

2   network of nodes each having a processor coupled to a memory, comprising:

3   representing the data contents of a memory of a node as a compressed Bloom filter data

4   structure stored in memory of a node having a preselected number of hash functions and a

5   preselected array size which have been chosen to optimize at least one of the rate of false positives

6   of the Bloom filter representing the data contents and the computational requirements of the

7   preselected number of hash functions for a transmission compression size;

8   compressing the Bloom filter data structure to the transmission compression size; and

9   periodically transmitting the compressed Bloom filter data structure to at least one other

10   node.

1   11. A method of storing data in memory for transmission, comprising:

2   representing the data as a compressed Bloom filter data structure in said memory having a

3   preselected number of hash functions and a preselected array size which have been chosen for a

4   target transmission compression size to optimize at least one of the rate of false positives of the

5   Bloom filter representing the data and the computational requirements of the preselected number of

6   hash functions.

1   12. A distributed computer network, comprising:

2   a plurality of periodically intercommunicating distributed network nodes;

3   each node including a cache memory and a processor coupled to the cache memory operative

4   to (1) represent its memory contents as a compressed Bloom filter data structure having a preselected

24

5   number of hash functions and a preselected array size which have been chosen for a target rate of

6   false positives to optimize at least one of the compression size of the Bloom filter representing the

7   memory contents and the computational requirements of the preselected number of hash functions,

8   to (2) compress the Bloom filter data structure to the target compression size using a predetermined

9   compression algorithm, and to (3) broadcast the compressed Bloom filter data structure to at least

10   one other node whenever the contents of its cache memory has changed.

1   13. The distributed computer network of claim 12, wherein said nodes are Web proxy servers.

1   14. The distributed computer network of claim 12, wherein said nodes are mobile or stationary

2   agents in a network of mobile nodes, and the Web objects correspond to agent locations.

1   15. The distributed computer network of claim 12, wherein said predetermined compression

2   algorithm is arithmetic coding.

1   16. A method employing compressed Bloom filters for storing and transmitting data in a distributed

2   network of nodes each having a processor coupled to a memory, comprising:

3   representing the data contents of a memory of a node as a compressed Bloom filter data

4   structure stored in memory of a node having a preselected number of hash functions and a

5   preselected array size which have been chosen to optimize at least one of the transmission

6   compression size of the Bloom filter representing the data contents and the computational

7   requirements of the preselected number of hash functions for a given rate of false positives;

8   compressing the Bloom filter data structure to the transmission compression size; and

9   periodically transmitting the compressed Bloom filter data structure to at least one other

10   node.

1   17. A method of storing data in memory for transmission, comprising:

2   representing the data as a compressed Bloom filter data structure in said memory having a

3   preselected number of hash functions and a preselected array size which have been chosen for a

4   target rate of false positives to optimize at least one of the transmission compression size of the

5   Bloom filter representing the data and the computational requirements of the preselected number of

6   hash functions.